

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Dempster-Shafer fusion of multisensor signals in nonstationary Markovian context

EURASIP Journal on Advances in Signal Processing 2012,
2012:134 doi:10.1186/1687-6180-2012-134

Mohamed El Yazid Boudaren (mohamed.boudaren@it-sudparis.eu)
Emmanuel Monfrini (Emmanuel.Monfrini@it-sudparis.eu)
Wojciech Pieczynski (Wojciech.Pieczynski@it-sudparis.eu)
Amar Aïssani (amraissani@yahoo.fr)

ISSN 1687-6180

Article type Research

Submission date 30 November 2011

Acceptance date 1 June 2012

Publication date 9 July 2012

Article URL <http://asp.eurasipjournals.com/content/2012/1/134>

This peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

For information about publishing your research in *EURASIP Journal on Advances in Signal Processing* go to

<http://asp.eurasipjournals.com/authors/instructions/>

For information about other SpringerOpen publications go to

<http://www.springeropen.com>

Dempster–Shafer fusion of multisensor signals in nonstationary Markovian context

Mohamed El Yazid Boudaren^{1,2,3,*}

* Corresponding author

Email: mohamed.boudaren@it-sudparis.eu

Emmanuel Monfrini²

Email: emmanuel.monfrini@it-sudparis.eu

Wojciech Pieczynski²

Email: wojciech.pieczynski@it-sudparis.eu

Amar Aïssani³

Email: aaissani@hotmail.com

¹ Ecole Militaire Polytechnique, Laboratoire Mathématiques Appliquées, P.O. Box 17, Bordj El Bahri, Algiers 16111, Algeria

² Institut Mines-Telecom, Telecom SudParis, CITI Dpt., CNRS UMR 5157, 9 rue Charles Fourier, Evry 91000, France

³ Université des Sciences et de la Technologie Houari Boumediene, LRIA, P.O. Box 32, Algiers 16111, Algeria

* Corresponding author. Université des Sciences et de la Technologie Houari Boumediene, LRIA, P.O. Box 32, Algiers 16111, Algeria

Abstract

The latest developments in Markov models' theory and their corresponding computational techniques have opened new rooms for image and signal modeling. In particular, the use of Dempster–Shafer theory of evidence within Markov models has brought some keys to several challenging difficulties that the conventional hidden Markov models cannot handle. These difficulties are concerned mainly with two situations: multisensor data, where the use of the Dempster–Shafer fusion is unworkable; and nonstationary data, due to the mismatch between the estimated stationary model and the actual data. For each of the two situations, the Dempster–Shafer combination rule has been applied, thanks to the triplet Markov models' formalism, to overcome the drawbacks of the standard Bayesian models. However, so far, both situations have not been considered in the same time. In this article, we propose an evidential Markov chain that uses the Dempster–Shafer combination rule to bring the effect of contextual information into segmentation of multisensor nonstationary data. We also provide the Expectation–Maximization parameters' estimation and the maximum posterior marginal's restoration procedures. To validate the proposed model, experiments are conducted on some synthetic multisensor data and noised images. The obtained segmentation results are then compared to those obtained with conventional approaches to bring out the efficiency of the present model.

Keywords

Hidden Markov chains, Triplet markov chains, Multisensor Markov chains, Data fusion, Theory of evidence, Dempster–Shafer combination rule, Multisensor data, Nonstationary data

Introduction

Hidden Markov chains (HMCs) have been used to solve a wide range of inverse problems occurring in many application fields. They allow one to take contextual information within data into account. Their success is mainly due to the existence of efficient Bayesian techniques that allow achieving the different estimation procedures with reasonable computational complexity. Hence, HMCs have successfully been applied in signal and image processing [1-3], biosciences [4], econometrics and finance [5], ecology [6], and communications [3]. Let us also mention [7-9] as pioneering articles.

Let $X = X_{1..N}$ be an unobservable process that takes its values from a finite set of classes $\Omega = \omega_1, \dots, \omega_K$ and let $Y = Y_{1..N}$ be an observable process that takes its values in \mathbb{R} and that can be seen as a noisy version of X . The problem is then to estimate X from Y . According to the HMC formalism, the hidden process X has a Markov distribution, and this is why the model is qualified by “hidden Markov”. The interest of considering X Markovian relies in the possibility of embedding the dependencies existing within Y into X . The observations Y_n are then assumed to be independent conditionally on X and the contextual information are considered only through X , which provides a well-designed formalism that permits to consider the data contextual information while keeping the model simple and the necessary estimation procedures workable. Explicitly, according to HMCs, the joint distribution of (X, Y) is given by

$$p(x, y) = p(x_1)p(x_1)p(y_1|x_1)\prod_{n=2}^N p(x_n|x_{n-1})p(y_n|x_n) \quad (1)$$

To estimate the hidden process of interest, one may use the Bayesian maximum posterior marginal’s (MPM) estimator that allows to minimize the ratio of erroneously assigned sites and that is given by

$$\hat{x}_n = \arg \max_{\omega \in \Omega} p(x_n = \omega | y) \quad (2)$$

The posterior distributions $p(x_n = \omega | y)$ are computable, thanks to the recursive Forward $a_n(x_n) = p(y_{1..n}, x_n)$ and Backward $\beta_n(x_n) = p(y_{n+1..N} | x_n)$ probabilities that can be computed iteratively as follows

$$\begin{cases} a_1(x_1) = p(x_1, y_1) \\ a_n(x_n) = \sum_{\omega \in \Omega} a_{n-1}(x_{n-1} = \omega) p(x_n | x_{n-1} = \omega) p(y_n | x_n) \end{cases} \quad (3)$$

$$\begin{cases} \beta_N(x_N) = 1 \\ \beta_n(x_n) = \sum_{\omega \in \Omega} \beta_{n+1}(x_{n+1} = \omega) p(x_{n+1} = \omega | x_n) p(y_{n+1} = \omega) \end{cases} \quad (4)$$

The estimator in Equation (2) can then be derived

$$\hat{x}_n = \arg \max_{\omega \in \Omega} a_n(x_n = \omega) \beta_n(x_n = \omega) \quad (5)$$

Notice that each \hat{x}_n is estimated using the whole observation $y_{1:N}$ and here relies the interest of the HMCs: they establish a link between all the variables $Y_{1:N}, X_{1:N}$ in such a way that each \hat{x}_n is estimated using all the $y_{1:n}$ while keeping the computation linear with the data size N . Moreover, when the model parameters are unknown, they can be estimated, thanks to some algorithms such as Expectation–Maximization (EM) [10] and Iterative Conditional Estimation (ICE) [11].

However, HMCs may become unworkable when the data to be modeled come from many heterogeneous sensors. Indeed, the conventional approaches involving Dempster–Shafer fusion (DS fusion) do not support Markov models, since such a fusion destroys Markovianity. Furthermore, standard HMCs have been shown to be inefficient when applied to nonstationary data when unsupervised processing is concerned. Let us mention, for instance, the situation when the distributions $p(x_{n+1} = \omega | x_n = \omega)$ depend on n .

Example 1: Let us consider the problem of satellite or airborne optical image segmentation into two classes $\Omega = \{\omega_1, \omega_2\}$ where ω_1 is “forest” and ω_2 is “water”. Let $(1, K, N)$ be the set of a line pixels of such an image. The problem is then to estimate the class-map $x_{1:N}$ given the observed image line $y_{1:N}$. The link between a pixel observation and its corresponding class is given by the likelihood probability $p(y_n | x_n = \omega)$, and the prior knowledge is modeled by a transition distribution $p(x_{n+1} = \omega | x_n = \omega')$. MPM estimation of $x_{1:N}$ is then workable according to Equation (5). Let us assume now that $p(x_{n+1} = \omega | x_n = \omega')$ depends on n . The use of standard HMCs in unsupervised segmentation in such a situation (nonstationary hidden process) provides poor results [12]. This is due to the mismatch between the estimated stationary model and the data.

The Dempster–Shafer theory of evidence [12–23] overcomes these drawbacks; thanks to the rich triplet Markov chains’ (TMC) formalism. In fact, the computation of posterior distribution $p(x|y)$, crucial for Bayesian restoration, can be seen as the DS fusion of the prior knowledge given by $p(x) = p(x_1) \prod_{n=2}^N p(x_n | x_{n-1})$ with the observation knowledge given by $q(x) \propto p(y|x) = \prod_{n=1}^N p(y_n | x_n)$. The result of such a fusion being linked with a TMC, the estimation algorithms remain workable.

Let us suppose now that we deal with more than one sensor, and that there are some clouds in the image provided by one of the sensors. The possible presence of clouds can be modeled by a probability measure on $P(\Omega) = \{\phi, \omega_1, \omega_2, \Omega\}$, which is a mass function [24].

The theory of evidence can then be utilized in the following situations where the use of conventional HMCs poses difficulties:

(1) When the prior distribution $p(x)$ is not known with precision (for instance, when

$p(x_n | x_{n-1})$ depends on n), it can be replaced by a belief function obtained from $p(x)$ to model the uncertainty or lack of accurate knowledge of $p(x)$. It can then be merged with $q(x)$ defined above via DS fusion. The result of this fusion gives a Bayesian probability on Ω^n that can be seen as a generalization of the posterior probability. Even if this latter is not necessarily Markovian, it is a marginal of a Markov chain and, thus, a Bayesian restoration remains workable [14].

(2) When the prior distribution is known with precision but one of the sensors is very noisy and its probabilistic noise densities are unreliable, or when both of the prior and noise distributions are exactly known, but there is an extra class in the data provided by one of the sensors.

Let us mention some previous studies that tackled the problem of using the theory of evidence in the Markovian context. The authors in [12] use evidential priors to deal with a strongly nonstationary prior distribution. The mass function extends then the Bayesian priors and the MPM restoration remains feasible; thanks to the TMC formalism. The resulting model is called evidential hidden Markov chain (EHMC). In [24], DS fusion is achieved in hidden Markov fields' context to merge images from more than one sensor, with some unreliable one. The aim of this article is to extend the previous results of the application of theory of evidence in Markovian context to the situations when $p(x)$ is nonstationary and when one of the sensors is unreliable at the same time. Hence, the use of the DS rule has two purposes: on the one hand, nonstationary data are modeled through an evidential model that considers uncertainty of the data distribution via evidential priors. On the other hand, the sensors' data are fused in Markovian context to improve the segmentation accuracy.

The remainder of the article is organized as follows: The following section summarizes the pairwise and triplet Markov chains formalisms. Section "Markov models and Dempster–Shafer theory of evidence" deals with the new trends in using the Dempster–Shafer theory of evidence in Markovian context. In Section "Multisensor nonstationary hidden Markov chains", we define the multisensor nonstationary HMC model and provide its corresponding MPM restoration and parameters estimation procedures. Experimental results are presented and discussed in Section "Experimental results". Finally, concluding remarks and some possible future improvements end the article.

Pairwise and triplet Markov chains

In this section, we briefly describe the pairwise Markov chains (PMCs) and the TMCs that are more general than the conventional HMCs defined in the previous section. In fact, PMCs in which the hidden process is not Markovian exist and are therefore more general. Similarly, TMCs form a family which is strictly more general than PMCs since TMCs which are not PMCs exist and have been used to deal with numerous situations that neither HMCs nor PMCs can support [25].

PMCs

Let $Z = (X, Y)$. Z is said to be a PMC if Z is itself Markovian. Therefore, Z is a PMC if and only if its joint distribution is given by

$$p(z) = p(z_1) \prod_{n=2}^N p(z_n | z_{n-1}) \quad (6)$$

An HMC defined by (1) can then be seen as a particular PMC in which $p(z_n | z_{n-1}) = p(x_n | x_{n-1})p(y_n | x_n)$, while in more general PMC, such a probability is given by $p(z_n | z_{n-1}) = p(x_n | x_{n-1}, y_{n-1})p(y_n | x_{n-1}, y_{n-1}, x_n)$. This shows the greater generality of PMC over

HMC at the local level. At the global level, on the other hand, the noise distribution $p(y|x)$ is of Markovian form in PMC whereas it is given by the simple formula $p(y|x) = \prod_{n=1}^N p(y_n | x_n)$ in conventional HMC. The posterior margins $p(x_n | y)$, needed for MPM restoration, are computable like in HMC model; thanks to the same forward functions $a_n(x_n) = p(y_1, K, y_1, x_n)$ and the extended backward functions that can be evaluated recursively as follows

$$\begin{cases} a_1(x_1) = p(x_1, y_1) \\ a_n(x_n) = \sum_{\omega \in \Omega} a_{n-1}(x_{n-1} = \omega) p(z_n | x_{n-1} = \omega, y_{n-1}) \end{cases} \quad (7)$$

$$\begin{cases} \beta_N(x_N) = 1 \\ \beta_n(x_n) = \sum_{\omega \in \Omega} \beta_{n+1}(x_{n+1} = \omega) p(x_{n+1} = \omega, y_{n+1} | z_n) \end{cases} \quad (8)$$

Besides, when the model parameters are unknown, they can be estimated via adapted variants of the same algorithms used for HMCs. For further details, the reader may refer to [26] where detailed related theoretical developments and experiments are provided.

TMCs

Z is referred to as a TMC if there exists a third process $U = U_{1:N}$ where each U_n takes its values from a finite set $\Lambda = \{\lambda_1, K, \lambda_M\}$ such that the triplet $T = (X, Y, U)$ is a Markov chain. Let $V = (U, X)$. $T = (V, Y)$ is then a PMC. This makes the computation of the distributions $p(x_n | y)$, required to perform MPM restoration, achievable even when Z is not Markovian. This shows the greater generality of TMC over PMC, which is itself more general than the conventional HMC [25].

The underlying process U may be used in all the situations where Z is a marginal distribution of a Markov chain. For instance, U is used to model the switches of the hidden process X [25], which constitutes in some manner, a way to deal with nonstationary aspect of X discussed in the previous section. The resulting model is called “switching hidden Markov chain”. Similarly, U has been used to consider the switches of the noise distributions in [27] and the semi-Markovianity of the hidden process in [28]. Another significant use of U is the

one used within the Dempster–Shafer theory of evidence to permit the use of this latter in the Markovian context as described in the following section.

Markov models and Dempster–Shafer theory of evidence

In this section, we briefly present the so-called theory of evidence introduced by Dempster in the 1960s and reformulated by Shafer in the 1970s [13]. Let $\Omega = \{\omega_1, \dots, \omega_K\}$ be a frame of discernment containing K exclusives hypotheses and let us consider the set of all the subsets of this frame $P(\Omega) = \{\emptyset, \{\omega_1\}, \dots, \Omega\}$. To understand the basics of the theory of evidence and to establish a link with the aim of the present framework, let us position ourselves in the satellite or airborne optical image segmentation problem of *Example 1*. The frame of discernment corresponds then to the set of hidden classes $\Omega = \{\omega_1, \omega_2\}$ where ω_1 and ω_2 correspond to “forest” and “water”, respectively. The exclusive hypotheses model the fact that one pixel of the image belongs either to the class “forest” or “water”. By considering compound hypotheses, the Dempster–Shafer theory of evidence offers an elegant formalism to model the uncertainty, the lack of precision or even missing information about the pixels classes using the so-called mass function. A mass function m is a function from $P(\Omega)$ to \mathbb{R}^+ that fulfills the following conditions

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{A \in P(\Omega)} m(A) = 1 \end{cases} \quad (9)$$

Notice that when the mass function m vanishes outside singletons, it becomes a probability, also called “Bayesian” or “probabilistic” mass in contrast to “evidential” mass according to the theory of evidence. Hence, it can be considered as a generalization of the probability measure. This generalization is the key notion that will be used to extend the conventional Bayesian models.

In the satellite segmentation problem, let us consider a pixelwise classification, the prior knowledge of the classes ω_1 and ω_2 may then be modeled by a Bayesian mass m_1 defined on Ω by $m_1(x_n = \omega_k) = p(x_n = \omega_k)$. On the other hand, the observation knowledge can be modeled through a Bayesian mass m_2 defined on Ω by $m_2(x_n = \omega_k) \propto p(y_n | x_n = \omega_k)$. The DS fusion gives $m(x_n) = m_1 \oplus m_2(x_n) = p(x_n | y_n)$. Suppose now that there are some clouds in the image provided by the sensor [24,29]. We have then three observable classes: “forest”, “water”, and “clouds”. The concept of evidential mass may be introduced here to model the fact that we cannot see through clouds. Explicitly, we may use a mass function m_2 defined on $\{\omega_1, \omega_2, \Omega\}$ by $m_2(\omega_1) \propto p(y_n | \omega_1)$, $m_2(\omega_2) \propto p(y_n | \omega_2)$, and $m_2(\Omega) \propto p(y_n | \Omega)$. Similarly, the DS fusion of the two masses $m(x_n) = m_1 \oplus m_2(x_n)$ gives a Bayesian mass that generalizes the posterior probability $p(x_n | y_n)$, where the DS combination rule of a set of mass functions m_1, \dots, m_r is given by the following formula

$$m(A) = (m_1 \oplus \dots \oplus m_r)(A) \propto \sum_{\substack{B_i \in P(\Omega) \\ B_i \cap A \neq \emptyset}} \prod_{i=1}^r m_i(B_i) \quad (10)$$

Let us bring up the following intuitive result: when one of the mass functions is Bayesian (probabilistic), the DS fusion result given by Equation (10) is also Bayesian [14].

Let us come back to the satellite or airborne optical image segmentation and let us assume now that the prior distribution of X is of Markovian form. The prior knowledge is then given by a Bayesian mass m_1 defined on Ω as follows: $m_1(x) \propto p(x_1)p(x_2|x_1)\mathbf{K} p(x_N|x_{N-1})$. On the other hand, the observation knowledge is modeled through a Bayesian mass m_2 defined on Ω as follows: $m_2(x) \propto p(y_1|x_1)\mathbf{K} p(y_N|x_N)$. The interesting result is that the DS fusion of the two masses is the posterior probability $p(x|y)$.

$$m(x) = (m_1 \oplus m_2)(x) = \frac{p(x)p(y|x)}{\sum_{x' \in \Omega^N} p(x')p(y|x')} = p(x|y) \quad (11)$$

The next step is then to take advantage of both Markov theory and theory of evidence by generalizing the Bayesian masses to evidential ones and exploit the result presented above. However, when at least one of the masses involved in the DS fusion is evidential, the result of this latter may no longer be a Markov chain, and thus, the Bayesian restoration is not directly applicable. The recent TMCs surmount this difficulty through the introduction of the third underlying process U as stated in the previous section. In fact, it has been shown that the DS fusion of the masses' functions defined above is a TMC and the calculation of the posterior distributions, necessary to achieve the different estimation procedures, remains possible.

As mentioned before, this result was used in [12] to consider evidential priors to take into account the nonstationary aspect of $p(x)$. Therefore, the authors define an “evidential HMC model” that extends the standard HMCs to the monosensor nonstationary case. On the other hand, the authors in [24] define a multisensor hidden Markov field that can resolve the problem of the image segmentation in presence of clouds discussed in this section. The aim of this article is to consider the problem where we have both situations simultaneously. Hence, the Dempster–Shafer theory of evidence is used, for one hand, to model the lack of precision in the prior distribution and, on the other hand, to consider more than one sensor with some unreliable one like the sensor with cloudy image.

Multisensor nonstationary hidden Markov chains

In this section, we describe our new model that will be called the multisensor nonstationary hidden Markov chain (MN-HMC), and we give its corresponding MPM restoration and EM parameters estimation procedures.

Model definition

Let $X = X_{1:K N}$ be a hidden process that takes its values from a finite set of classes $\Omega = \{\omega_1, \mathbf{K}, \omega_K\}$ and that is to be estimated from a family of observable processes $Y = Y^{1:K R}$ provided by R independent sensors $S_{1 \dots R}$ where $Y^r = Y_{1 \dots N}^r$ and where each Y_n^r takes its values from \mathbb{R} . Let us assume now that the realization of X is governed by a nonstationary Markov

chain in the sense that the distributions $P(x_{n+1} = \omega | x_n = \omega')$ are not known with precision (or depend on n). Let us also suppose that at least one of the sensors $S_{1...R}$ is Bayesian. Without loss of generality, let S_1 be such a sensor. The MPM Bayesian restoration using the whole observable process $Y = Y^{\text{IK}^R}$ is then workable and the number of elementary operations required for its evaluation is linear with the size of the data N (the proof can be found in [14]). The model $Z = (X, Y)$ is then called an MN-HMC. The MPM restoration of the hidden process of an MN-HMC can then be achieved; thanks to the DS fusion of the different mass functions involved in the model.

$$m(x) = (m_0 \oplus L \oplus m_R)(x) = p(x|y) \quad (12)$$

Although the result of this fusion is not necessarily a Markov chain, it is a marginal of a TMC [14] and hence, the posterior marginal distributions $P(x_n|y)$ are computable.

Let us now demonstrate how the data are modeled according to this model.

First, the nonstationary Markov chain governing the hidden process X is replaced by a stationary evidential Markov chain m_0 . Let $U^0 = U_{\text{IK}^N}^0$ be a hidden process that takes its values in $P(\Omega) = \{\phi, \{\omega_1\}, K, \Omega\}$, m_0 is then defined on $[P(\Omega)]^N$.

Second, the sensor S_1 being Bayesian, its mass m_1 is then defined on Ω as follows: $m_1(x) \propto p(y_1^1|x_1)p(y_2^1|x_2) \dots p(y_N^1|x_N)$.

For each sensor, we derive the corresponding observation mass. For example, let us suppose that sensor S_2 is only sensitive to class ω_1 . The corresponding evidential mass is then defined on $[\Lambda_2]^N$ with $\Lambda_2 = \{\omega_1\}, \{\omega_2, K, \omega_K\}$. For this purpose, we consider an underlying process $U^2 = U_{\text{IK}^N}^2$ that takes its values in Λ_2 .

Finally, for some sensors, we may consider more than one mass function. Let us consider the sensor of *Example 1* and let us assume that there are some clouds in the image that it provides. The possible presence of clouds can then be modeled by a probability measure on $P(\Omega) = \{\phi, \omega_1, \omega_2, \Omega\}$, which is a mass function. More explicitly, we consider an underlying process $U^2 = U_{\text{IK}^N}^3$ that takes its values in $\{\omega_1, \omega_2, \Omega\}$. Furthermore, we can consider an additional mass function to model the prior knowledge of cloud presence regardless of the sensor observation. This shows again the greater generality of evidential models over Bayesian ones.

Let $M = m_{\text{OK}^S}$ be the set of masses that model the data under consideration. For one hand, we know that the DS fusion result of all these mass functions is a TMC, and on the other hand, the result of this fusion is the posterior probability $p(x|y)$. Hence, the posterior distributions $P(x_n|y)$ necessary to achieve MPM restoration are computable.

Unsupervised segmentation of MN-HMCs

Let us consider the following image segmentation problem that extends the one given in *Example 1*.

Example 2: Let us consider the problem of satellite or airborne optical image segmentation into two classes $\Omega = \{\omega_1, \omega_2\}$ where ω_1 corresponds to “forest” and ω_2 to “water”. Let S_1 and S_2 be two sensors, where S_1 is a RADAR sensor and S_2 is an optical one. Let $(1, \dots, N)$ be the set of a line sites of the ground truth image X . The problem is then to estimate the class-map $x_{1:N}$ given the observations $y_{1:N}^1$ and $y_{1:N}^2$ provided by S_1 and S_2 , respectively.

Sensor 1: The digital image observations $y_{1:N}^1$ are related to the hidden classes through the noise probability density functions (pdfs) $p(y_n^1 | x_n = \omega_1)$ and $p(y_n^1 | x_n = \omega_2)$.

Sensor 2: Let us assume that there are some clouds in the image $y_{1:N}^2$ provided by S_2 . We have then three possibilities: “forest”, “water”, and “clouds”. The observation at each pixel is related to its class through the noise pdfs $p(y_n^2 | x_n = \omega_1)$, $p(y_n^2 | x_n = \omega_2)$, and $p(y_n^2 | x_n = \text{"clouds"})$, respectively.

The problem consists then in how to estimate the class-map $x_{1:N}$ using both sensors images in Markovian context in such a way that nonstationary aspect of $p(x)$ is taken into account. For this purpose, we use our proposed model. First, we have to gather all the information that can be fused lately to achieve the MPM restoration.

Data modeling

First, the nonstationary Markov chain governing X is replaced by a stationary evidential Markov chain m_0 . Let $U^1 = U_{1:N}^1$ be a hidden process that takes its values in $P(\Omega) = \{\phi, \{\omega_1\}, \{\omega_2\}, \Omega\}$ and that models the lack of knowledge of X priors, m_0 is then defined on $[P(\Omega)]^N$ by $m_0(u^1) = m_0(u_1^1) \prod_{n=2}^N m_0(u_n^1 | u_{n-1}^1)$.

The first sensor being Bayesian, the observation knowledge may then be modeled through a probabilistic mass function m_1 defined by $m_1(x) = \prod_{n=1}^N m_1(x_n) \propto \prod_{n=1}^N p(y_n^1 | x_n)$. Accordingly, the MPM restoration may be done based on this sensor alone.

The possible presence of clouds may be modeled by a probability measure on $P(\Omega) = \{\phi, \omega_1, \omega_2, \Omega\}$ which is a mass function given by $m_2(u^2) \propto \prod_{n=1}^N m_2(u_n^2)$ where $u^2 \in [P(\Omega)]^N$ and $m^2(\omega_1) \propto p(y_n^2 | u_n^2 = \omega_1)$, $m_2(\omega_2) \propto p(y_n^2 | u_n^2 = \omega_2)$, and $m_2(\Omega) \propto p(y_n^2 | u_n^2 = \Omega)$. For this particular sensor, we may use another evidential Markov mass to model the contextual information corresponding to the presence of clouds. In fact,

neighboring clouds' pixels belong more likely to cloud class than other pixels do. Let m_3 be such a mass defined on $[P(\Omega)]^N$ by $m_3(u^2) = m_3(u_1^2) \prod_{n=2}^N m_3(u_n^2 | u_{n-1}^2)$.

MPM restoration of hidden data

We have defined a family of four masses m_0, m_1, m_2 , and m_3 that represent all the knowledge we have about the data. Let $T = (X, U, Y) = (V, Y)$ be a TMC where $X = X_{1:N}$ with each $x_n \in \Omega$, $u_n \in [P(\Omega)]^2$ with each $u_n \in [P(\Omega)]^2$ and $Y = (Y^1, Y^2)$ with each $y_n \in \mathbb{R}^2$. The distribution of T is then given by

$$p(t_1, K, t_n) \propto q_1(t_1, t_2) q_2(t_2, t_3) \dots q_{N-1}(t_{n-1}, t_N) \quad (13)$$

where

$$\begin{cases} q_1(t_1, t_2) = 1_{[x_1 \in u_1^1 | u_1^2]} 1_{[x_2 \in u_2^1 | u_2^2]} \times \\ m_0(u_1^1) m_0(u_2^1 | u_1^1) m_1(x_1) m_1(x_2) m_2(u_1^2) m_2(u_2^2) m_3(u_1^2) m_3(u_2^2 | u_1^2) \\ q_n(t_n, t_{n+1}) = 1_{[x_{n+1} \in u_{n+1}^1 | u_{n+1}^2]} m_0(u_{n+1}^1 | u_n^1) m_1(x_{n+1}) m_2(u_{n+1}^2) m_3(u_{n+1}^2 | u_n^2) \end{cases} \quad (14)$$

Then, the DS fusion result $m = (m_0 \oplus m_1 \oplus m_2 \oplus m_3)$ is the posterior distribution $p(x|y)$ defined by the joint distribution $p(x_n, u_n | y)$ which is itself the marginal distribution of the TMC defined above. Hence, the posterior marginal distributions $p(x_n | y)$ are computable and so are the interesting probabilities $p(x_n | y)$.

Finally, to achieve the MPM restoration of the hidden process, we either use the theorem giving a general definition of a Markov chain [12] or the well-known forward $a_n(v_n) = p(y_1, K, y_n, v_n)$ and backward $\beta_n(v_n) = p(y_{n+1}, K, y_N | v_n)$ recursive functions that have been adapted to the multisensor nonstationary context. In this article, we chose to adopt the latter option. For the use of the Markov chain theorem, the reader may refer to [12].

The forward and backward functions can be calculated in the following iterative ways:

$$\begin{cases} a_1(v_1) = p(v_1, y_1) \\ a_n(v_n) = \sum_{v_{n-1}} a_{n-1}(v_{n-1}) p(v_n | v_{n-1}) m_1(x_n) m_2(u_n^2) \end{cases} \quad (15)$$

$$\begin{cases} \beta_N(x_N) = 1 \\ \beta_n(v_n) = \sum_{v_{n+1}} \beta_{n+1}(v_{n+1}) p(v_{n+1} | v_n) m_1(x_{n+1}) m_2(u_{n+1}^2) \end{cases} \quad (16)$$

Where $p(v_{n+1} | v_n) \propto 1_{[x_n \in u_n^1 | u_n^2]} 1_{[x_{n+1} \in u_{n+1}^1 | u_{n+1}^2]} m_0(u_{n+1}^1 | u_n^1) m_3(u_{n+1}^2 | u_n^2)$ and $p(v_1, y_1) \propto 1_{[x_1 \in u_1^1 | u_1^2]} m_0(u_1^1) m_1(x_1) m_2(u_1^2) m_3(u_1^2)$.

The posterior marginal distributions $p(v_n|y)$ and $p(x_n|y)$ can then be computed as follows

$$p(v_n|y) = \alpha_n(v_n)\beta_n(v_n) \quad (17)$$

$$p(x_n|y) = \sum_{v_n \ni x_n} p(v_n|y) \quad (18)$$

On the other hand, the posterior transitions and marginal distributions necessary for the parameters estimation can be calculated according to

$$\psi(v_n, v_{n+1}) = p(v_n, v_{n+1}|y) = \alpha_n(v_n)p(v_n|v_{n-1})m_1(x_{n+1})m_2(u_{n+1}^2)\beta_n(v_{n+1}) \quad (19)$$

$$\xi(v_n) = p(v_n|y) = \sum_{v_{n-1}} \psi(v_{n-1}, v_n) \quad (20)$$

Model parameters estimation

To estimate the model parameters, we either use the well-known EM algorithm, its stochastic version SEM or the ICE algorithm. Let us mention that all these latter have been used in the triplet Markov models context [25,28]. Let us also mention [30] where a brief comparative study is conducted between EM and ICE algorithms in the hidden, pairwise, and triplet Markov models contexts. As we deal with a particular TMC model, we only need to adapt each one to the situation addressed in this article.

In this article, we propose to adapt the EM algorithm to the MN-HMC case. For this purpose, let us consider the TMC $T = (X, U, Y) = (V, Y)$ defined above. For the sake of simplicity, we

consider the Gaussian case where the noise *pdfs* $p(y_n|x_n)$ and $p(y_n|u_n^2)$ are of Gaussian form. According to *Example 2* (where $\Omega = \{\omega_1, \omega_2\}$), we have to estimate the following parameters: the evidential mass $m_{ij} = m_0(u_n^1 = \lambda_i, u_{n+1}^1 = \lambda_j)$ defined on $[P(\Omega)]^2$, the $K = 2$ means $\mu_{1..K}^1$ standard deviations of the Gaussian *pdfs* governing m_1 , the $K + 1 = 3$ means $\mu_{1..K+1}^2$ and standard deviations $\sigma_{1..K+1}^1$ of the Gaussian densities governing m_2 and the evidential transition mass $c_{ij} = m_3(u_n^2 = \lambda_i | u_{n+1}^2 = \lambda_j)$ defined on $[P(\Omega)]^2$. The parameters estimation process is accomplished iteratively as follows:

(1) Consider an initial set of parameters $\Theta^0 = (m_{ij}^0, c_{ij}^0, (\mu_{1..K}^1, \mu_{1..K}^1, \sigma_{1..K}^2, \sigma_{1..K}^2)^0)$.

(2) For each iteration q , calculate Θ^{q+1} from Θ^q and Y in two steps:

- a. *Step E*: compute $\alpha_n^q(v_n)$ and $\beta_n^q(v_n)$ and then derive $\psi^q(v_n, v_{n+1})$ and $\xi^q(v_n)$.
- b. Compute Θ^{q+1} as follows.

$$(\mu_k^1)^{q+1} = \frac{\sum_{n=1}^N \sum_{u_n} \xi^q(x_n = \omega_k, u_n) y_n^1}{\sum_{n=1}^N \sum_{u_n} \xi^q(x_n = \omega_k, u_n)} \quad (21)$$

$$(\mu_k^2)^{q+1} = \frac{\sum_{n=1}^N \sum_{v_n/u_n^2=\lambda_k} \xi^q(v_n) y_n^2}{\sum_{n=1}^N \sum_{v_n/u_n^2=\lambda_k} \xi^q(v_n)} \quad (22)$$

$$[(\sigma_k^1)^{q+1}]^2 = \frac{\sum_{n=1}^N \sum_{u_n} \xi^q(x_n = \omega_k, u_n) y_n^1 - (\mu_k^1)^{q+1}^2}{\sum_{n=1}^N \sum_{u_n} \xi^q(x_n = \omega_k, u_n)} \quad (23)$$

$$[(\sigma_k^2)^{q+1}]^2 = \frac{\sum_{n=1}^N \sum_{v_n/u_n^2=\lambda_k} \xi^q(v_n) y_n^2 - (u_k^2)^{q+1}}{\sum_{n=1}^N \sum_{v_n/u_n^2=\lambda_k} \xi^q(v_n)} \quad (24)$$

$$m_{ij}^{q+1} = \frac{1}{(N-1)\#i\#j} \sum_{n=1}^N \sum_{v_n, v_{n+1}/u_n^1=\lambda_i, u_{n+1}^1=\lambda_j} \psi^q(v_n, v_{n+1}) \quad (25)$$

$$c_{ij}^{q+1} = \frac{\sum_{n=1}^N \sum_{v_n, v_{n+1}/u_n^2=\lambda_i, u_{n+1}^2=\lambda_j} \psi^q(v_n, v_{n+1})}{\sum_{n=1}^N \sum_{v_n/u_n^2=\lambda_i} \xi^q(v_n)} \quad (26)$$

where $\#i$ is the cardinal of the set λ_i .

Experimental results

This section is devoted to the application of the MN-HMC, described above, to the segmentation of multisensor nonstationary signals.

For this purpose, let us consider the following situation:

Let S_1 and S_2 be two sensors providing two different observable signals $y^1 = y_{\text{IK } N}^1$ and $y^2 = y_{\text{IK } N}^2$ that can be seen, in some manner, as noisy versions of a ground truth $x = x_{\text{IK } N}$ with the following difficulties: X , which is hidden and to be estimated in some way from $Y = (Y^1, Y^2)$, is a realization of an unknown Markov chain that may be strongly nonstationary. Moreover, the signal Y^2 presents an extra class which may correspond to clouds in a SPOT image or even missing observation at some signal sites. This class represents then the ignorance attached with the fact that we cannot decide whether such a site belongs to any of the classes. Let B be the process corresponding to the presence of this extra class. In this study, such a process is assumed to be Markovian.

Thereafter, we consider two series of experiments: in the first one, we deal with synthetic multisensor nonstationary data, whereas in the second one, we consider two nonstationary class-images that we noise in some manner to fit the multisensor nonstationary context. To assess the efficiency of the proposed model, MPM restoration is also achieved according to some conventional models.

Unsupervised segmentation of MN-HMCs

In this experiment, we deal with sampled multisensor nonstationary HMCs. Let $T = (X, U^2, Y)$ be such a model with $\Omega = \{\omega_1, \omega_2\}$ and $N = 5000$ and the following matrices

$$M_1 = \begin{bmatrix} 0.98 & 0.02 \\ 0.02 & 0.98 \end{bmatrix}, M_2 = \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{bmatrix}, \text{ and } J = \begin{bmatrix} 0.98 & 0.02 \\ 0.01 & 0.99 \end{bmatrix}.$$

The hidden process $X = X_1 \dots X_N$ is nonstationary in the following way: given the two matrices M_1, M_2 , and a value of $s = 1, 2, K$, $X^i = (X_{(i-1)s+1}, X_{(i-1)s+2}, \dots, X_{is})$. The realization of X fulfills the following

- The distribution of X_1 is $(0.5, 0.5)$.

- M_1 is the transition matrix in X^1, X^3, \dots
- M_2 is the transition matrix in X^2, X^4, \dots

A realization of X is then sampled. On the other hand, the realization of B , modeling the presence of the extra-class in the second sensor signal, is sampled as a Markov chain with transition matrix J . Accordingly, the corresponding realization of U^2 can be derived as follows

$$\begin{cases} u_n^2 = \Omega \text{ if } b_n = 1 \\ u_n^2 = x_n \text{ elsewhere} \end{cases} \quad (27)$$

Given the realizations of X and U^2 , the observed signals are then sampled in the following manner

- *Sensor 1:* $y^1 = y_{1KN}^1$ is sampled according to $p(y_1^1|x_1)p(y_2^1|x_2)\dots p(y_N^1|x_N)$ where $p(y_n^1|x_n = \omega_1)$ is Gaussian with mean 0 and standard deviation 1, and $p(y_n^1|x_n = \omega_2)$ is Gaussian with mean 1 and standard deviation 1.

- *Sensor 2:* $y^2 = y_{2KN}^2$ is sampled according to $p(y_1^2|u_1^2)p(y_2^2|u_2^2)\dots p(y_N^2|u_N^2)$ where $p(y_n^2|u_n^2 = \{\omega_1\})$ is Gaussian with mean 0 and standard deviation 1, $p(y_n^2|u_n^2 = \{\omega_2\})$ is Gaussian with mean 1 and standard deviation 1, and $p(y_n^2|u_n^2 = \{\omega_1, \omega_2\})$ is Gaussian with mean 2 and standard deviation 1.

As these experiments aim at assessing the proposed model against the conventional ones, MPM restoration of the hidden process of interest is also achieved according to the following family of approaches (with increasing degree of complexity):

- The segmentation is accomplished based on the first sensor signal using: K-Means, standard HMC, and evidential HMC.
- Then, we achieve MPM segmentation based on both sensors images using the multisensor stationary HMC (MS-HMC) as the one proposed in the Markov fields' context [24], and the MN-HMC formalism, proposed in this article.

Hundred experiments are carried out for each value of s . The obtained segmentation results are summarized in Table 1. From the obtained segmentation error ratios, we can mention the following remarks:

- The segmentation error ratios obtained by the application of K-Means and HMCs do not depend on the value of s .
- As the value of s is higher, as the data nonstationarity is stronger. Since we deal with unsupervised segmentation, the evidential HMC outperforms the conventional HMC. This is due to the fact that the evidential HMC takes into account the nonstationary aspect of the data. Similarly, the multisensor nonstationary HMC outperforms the multisensor stationary HMC for high values of s .
- Both multisensor models outperform the standard HMC. This difference is due to the fact that they utilize more data than the conventional HMC does. This also shows that the EM

procedure yields better parameters estimates in the unsupervised context when more data are available.

- The importance of considering data nonstationarity against the importance of the amount of data used to achieve the segmentation can be evaluated by comparing error ratios of both the EHMC with the MS-HMC. When the data are strongly nonstationary (high values of s), the EHMC provides better results than the MS-HMC since this latter does not take nonstationarity into account. On the other hand, for low values of s , the MS-HMC yields better results than the EHMC.

- Overall, the proposed MN-HMC outperforms the previous models. In fact, the proposed model utilizes more than one observed signal while it takes into account the nonstationary aspect of the hidden process. It benefits on one hand, from the advantages of the contextual information through the use of Markov theory and, on the other hand, from the benefits of the theory of evidence that permits to consider uncertainty in hidden classes priors and data fusion in the same time.

Table 1 Error ratios (%) of unsupervised segmentation of synthetic multisensor nonstationary data

s	$\tau_{\text{K-Means}}$	$\tau_{\text{HMC-EM}}^{\text{MPM}}$	$\tau_{\text{EHMC-EM}}^{\text{MPM}}$	$\tau_{\text{MS-HMC-EM}}^{\text{MPM}}$	$\tau_{\text{MN-HMC-EM}}^{\text{MPM}}$
2	30.6	26.7	26.7	22	22
4	32.3	25.7	25.6	19.7	19.7
8	31.5	23.6	24	20	19.7
16	30.3	22.5	22.3	19.2	17.7
32	31.6	24.5	22.2	20.4	17.5
64	30.9	23.1	21.6	19.7	17.2
128	30.7	22.4	19.6	18.4	15.5
256	29.9	25.7	19.2	19.1	15.4
512	32.4	23.5	18.2	21.5	15.3
1024	31.6	22.9	17.7	20.7	15.3
2048	31	25.3	17.1	17.9	14.1

Unsupervised segmentation of multisensor noisy nonstationary images

In this experiment, we propose to apply our model to multisensor noisy nonstationary images. To make our chain model applicable on images, these latter are converted from and to 1Dsignals using the Hilbert-Peano scan [31].

For this experiments set, we consider two nonstationary class images: the “Nazca bird” image (Figure 1a) and the “squares” image (Figure 2a).

Figure 1 Unsupervised segmentation of multisensor noisy image 1. (a) Original class image $X = x$. (b) First sensor observed image $Y^1 = y^1$. (c) Second sensor observed image $Y^2 = y^2$. (d) Image restoration using K-Means, error ratio $\tau_{\text{K-Means}} = 39.4\%$. (e) Image MPM restoration using HMCs, error ratio $\tau_{\text{HMC}} = 16.5\%$. (f) Image MPM restoration using evidential HMCs, error ratio $\tau_{\text{EHMC}} = 11.1\%$. (g) MPM estimation of the auxiliary process U^1 according to evidential HMCs. (h) Image MPM restoration using multisensor stationary HMCs, error ratio $\tau_{\text{MS-HMC}} = 12\%$. (i) Estimation of the underlying U^2 process according to

multisensor stationary HMCs, error ratio $\tau_{\text{MS-HMC}} = 0.1\%$. **(j)** Image MPM restoration using multisensor nonstationary HMCs, error ratio $\tau_{\text{MN-HMC}} = 7.9\%$. **(k)** MPM estimation of the auxiliary process U^1 according to multisensor nonstationary HMCs. **(l)** Estimation of the underlying U^2 process according to multisensor nonstationary HMCs, error ratio $\tau_{\text{MN-HMC}} = 0.1\%$

Figure 2 Unsupervised segmentation of multisensor noisy image 2. **(a)** Original class image $X = x$. **(b)** First sensor observed image $Y^1 = y^1$. **(c)** Second sensor observed image $Y^2 = y^2$. **(d)** Image restoration using K-Means, error ratio $\tau_{\text{K-Means}} = 30.9\%$. **(e)** Image MPM restoration using HMCs, error ratio $\tau_{\text{HMC}} = 16.4\%$. **(f)** Image MPM restoration using evidential HMCs, error ratio $\tau_{\text{EHMC}} = 15.9\%$. **(g)** MPM estimation of the auxiliary process U^1 according to evidential HMCs. **(h)** Image MPM restoration using multisensor stationary HMCs, error ratio $\tau_{\text{MS-HMC}} = 9.1\%$. **(i)** Estimation of the underlying U^2 process according to multisensor stationary HMCs, error ratio $\tau_{\text{MS-HMC}} \approx 0\%$. **(j)** Image MPM restoration using multisensor nonstationary HMCs, error ratio $\tau_{\text{MN-HMC}} = 7.6\%$. **(k)** MPM estimation of the auxiliary process U^1 according to multisensor nonstationary HMCs. **(l)** Estimation of the underlying U^2 process according to multisensor nonstationary HMCs, error ratio $\tau_{\text{MN-HMC}} \approx 0\%$

Let us consider, for instance, the “Nazca bird” nonstationary image which is a 128×128 class-image with $K=2$ classes that will serve as a ground truth image. Then, we noise the image in two different manners to have two observed images that can be lately fused using the proposed MN-HMC. Hence, we have $\Omega = \{\omega_1, \omega_2\}$ and $N = 16384$.

For the first observed image, $y^1 = y_{1:N}^1$ is sampled according to $p(y_1^1|x_1)p(y_2^1|x_2)\dots p(y_N^1|x_N)$ where $p(y_n^1|x_n = \omega_1)$ is Gaussian with mean 0 and standard deviation 1 and $p(y_n^1|x_n = \omega_2)$ is Gaussian with mean 1 and standard deviation 1.

For the second observed image, let us assume that some of the image pixels are corrupted (or even missing). We have then, three classes: ω_1 , ω_2 , and an extra-class where we cannot decide whether the given pixel belongs to either of the two classes. Let B be the process that governs the presence of the third class (that we call “*corrupted*”). In this experiment, we assume this latter to be Markovian. The realization of this latter was sampled according to the following transition matrix defined on the set $\{1, 2\}$ where ‘1’ corresponds to “*corrupted*” and ‘2’ corresponds to “*not corrupted*”.

$$J = \begin{bmatrix} 0.998 & 0.002 \\ 0.001 & 0.999 \end{bmatrix}.$$

Accordingly, the corresponding realization of U^2 can be derived as follows

$$\begin{cases} u_n^2 = \Omega \text{ if } b_n = 1 \\ u_n^2 = x_n \text{ elsewhere} \end{cases} \quad (28)$$

$y^2 = y_{\text{IK}N}^2$ is then sampled according to $p(y_1^2|u_1^2)p(y_2^2|u_2^2)K p(y_N^2|u_N^2)$ where $p(y_n^1|u_n^2 = \{\omega_1\})$ is Gaussian with mean 0 and standard deviation 1, $p(y_n^1|u_n^2 = \{\omega_2\})$ is Gaussian with mean 2 and standard deviation 1, and $p(y_n^1|u_n^2 = \{\omega_1, \omega_2\})$ is Gaussian with mean 4 and standard deviation 1.

Notice that the contrast between the two classes ω_1 and ω_2 is higher in the second image. However, the reliability of the corresponding sensor (presence of a third class) makes the direct application of conventional hidden Markov models unworkable. The same thing happens when the observation is missing in some pixels (which can be seen as a particular case of the present one). This challenging difficulty can be surmounted; thanks to the evidential model proposed in [24] and the one proposed in this framework.

The MPM restoration of the class-image is then achieved using the following approaches:

- The segmentation is carried out based on the first sensor signal only, using: K-Means, standard HMC, and EHMC.
- Then, we achieve MPM segmentation using both the MS-HMC and the MN-HMC.

The MPM segmentation results are shown in Figure 1.

When applying K-Means clustering algorithm, the only information used to restore data are the direct observations whereas no prior information about classes are considered. Consequently, this model is too sensitive to noise and the segmentation error ratio is relatively high $\tau_{\text{K-Means}} = 39.4\%$.

In conventional HMCs, the neighborhood of each site is taken into account. However, nonstationary aspect of the hidden data makes the restoration results poor $\tau_{\text{HMC}} = 16.5\%$. Indeed, as can be seen in the original class-image, the two classes are not distributed in the same manner along the image; there are some regions with a lot of details (wings and tail of the bird), whereas the image background is characterized by only one class (white). This particular aspect of the class-image has misled the segmentation through HMC since all their corresponding estimation procedures consider the hidden process X as stationary.

The application of the EHMC permits to overcome the difficulty discussed above ($\tau_{\text{EHMC}} = 11.1\%$) through the introduction of a mass function that generalizes the Bayesian prior probabilities and takes into account the uncertainty attached to the prior distribution of X , due to the heterogeneous aspect of the two classes along the signal. However, it would be interesting to make use of the second image where the contrast between the two classes is higher, even if some pixels are hidden with an extra-class.

Conversely, the MS-HMC exploits all the observed data, and provides then better results than the conventional HMC. Nevertheless, it does not take the nonstationary aspect of the data into account, and provides then comparable results with the evidential HMC $\tau_{\text{MS-HMC}} = 12\%$.

Finally, the MN-HMC yields the best result among all the considered models ($\tau_{\text{MN-HMC}} = 7.9\%$). This is due to the fact that this model takes advantages of two observed images while it takes the nonstationarity of the data into account. The evidential HMC can then be seen as a particular case of the MN-HMC, when only one sensor is available, whereas the

MS-HMC can be considered as a particular case of the MN-HMC where the data to be modeled are actually stationary.

In the image corresponding to the estimation of the process U^2 (Figure 1k), the region in white corresponds to the sub-set Ω where the confusion between the two classes is too high. In fact, this region of the image (that corresponds to the wings and tail of the bird) is characterized by a lot of details. Let us mention that for such a region, K-Means may provide comparable, and may be even, better segmentation results than standard HMC and MS-HMC. This is due to the fact that the regularization in both HMC and MS-HMC misleads the classification process in this region when considering $p(x)$ not depending on n . The aim of the use of the EHMC relies in weakening the prior knowledge about hidden classes in such regions to consider rather observation knowledge. In the same region of interest (wings and tail of the bird), our MN-HMC model provides also the best result because it uses more information than the other models do (two sensors rather than one) while it takes into account the nonstationarity of the data.

The EM-estimated parameters according to all the Markov models are also provided in Table

2. The real Gaussian noise *pdfs* parameters being known $\mu_1^1 = 0, \mu_2^1 = 1, \mu_1^2 = 0, \mu_2^2 = 2, \mu_{1,2}^2 = 4$, we can check that the EM-estimated parameters according to the proposed MN-HMC are closer to the real ones. Let us focus on the EM-estimated parameters according to the different considered Markov models:

- According to the standard HMC, the data provided by the first sensor are considered stationary. Hence, the HMC regularization misleads the parameters estimations process. The same thing happens in the MS-HMC context: even if both sensors images are used, there is a mismatch between the data and the EM-estimated stationary model, which leads to unsuitable estimated parameters set.

- The evidential HMC model takes into account the nonstationary aspect of data but only considers the image provided by the first sensor. Therefore, the estimated parameters are close to the real ones, but the MPM segmentation is based only on one image and hence the segmentation results are relatively limited.

- The parameters estimated based on the proposed model are the closest to the genuine ones. Besides, we can measure the difference between the parameters estimated according to the two multisensor models: since the MS-HMC is a particular MN-HMC where the transition mass vanishes outside the singletons $\{\omega_1\}$ and $\{\omega_2\}$, the relatively high probability 0.2540 attributed to $m(\{\omega_1, \omega_2\}, \{\omega_1, \omega_2\})$ by the proposed MN-HMC, can be seen as an inadequacy measure of the multisensor stationary HMC.

Table 2 EM-estimated parameters of the “Nazca bird” noisy multisensor image according to different Markov models

<i>HMC parameters estimated by EM</i>			
		μ	σ
Noise densities parameters	ω_1	0.3593	1.0331
	ω_2	1.1189	0.9526
Transition matrix		ω_1	ω_2
	ω_1	0.7328	0.2672
	ω_2	0.1525	0.8475
<i>Evidential HMC parameters estimated by EM</i>			

		μ		σ
Noise densities parameters				
	ω_1	-0.0652		0.9498
	ω_2	1.0284		0.9676
	ω_1	ω_2		Ω
Transition mass	ω_1	0.0105	0.0403	0.0008
	ω_2	0.0402	0.6726	0.0008
	Ω	0.0008	0.0008	0.2332
<i>Multisensor stationary HMC parameters estimated by EM</i>				
		μ^1	μ^2	σ^1
Noise densities parameters	ω_1	0.2964	0.6025	1.0745
	ω_2	1.0061	1.9891	0.9831
	Ω	—	4.0104	—
		ω_1		ω_2
Transition mass	ω_1	0.2081		0.0219
	ω_2	0.0219		0.7481
<i>Multisensor nonstationary HMC parameters estimated by EM</i>				
		μ^1	μ^2	σ^1
Noise densities parameters	ω_1	0.0046	-0.0356	0.9994
	ω_2	0.9928	1.9893	0.9843
	Ω	—	4.0099	—
		ω_1	ω_2	Ω
Transition mass	ω_1	0.0074	0.0108	0.0022
	ω_2	0.0106	0.7096	0.0016
	Ω	0.0024	0.0014	0.2540

The segmentation results of the “squares” nonstationary image confirm the previous comments with the following slight difference: the segmentation error ratio based on the evidential HMC model ($\tau_{\text{EHMC}} = 15.9\%$) is significantly higher than that of the multisensor stationary HMC ($\tau_{\text{MS-HMC}} = 9.1\%$). This is due to the fact that the considered “squares” image is moderately nonstationary. In such cases, the effect of the amount of data used to achieve the MPM segmentation is more important than the effect of taking nonstationarity into account. The same concluding remark has been mentioned in the synthetic experiments set for low values of s . Let us mention that, for such kind of data, the gain in segmentation accuracy by using the proposed model is also restrained ($\tau_{\text{MN-HMC}} = 7.6\%$ against $\tau_{\text{MS-HMC}} = 9.1\%$).

Conclusions

In this article, we proposed a new approach to model multisensor nonstationary signals in the Markovian context. The proposed model allows one to benefit simultaneously from both the Markov theory and theory of evidence. Accordingly, Dempster–Shafer combination rule was used for two purposes: to take into account the nonstationary aspect of the hidden data of interest, and to fuse the different sensors’ signals in the Markovian context to boost up the segmentation accuracy. The experimental results demonstrated the interest of such a modeling with respect to the conventional ones. As future improvement, we may investigate the use of evidential pairwise Markov models to consider more complex model structures. A

further generalization of the present approach may also consist in adapting the proposed formalism to Markov trees models in order to model multi-resolution images [32].

Competing interest

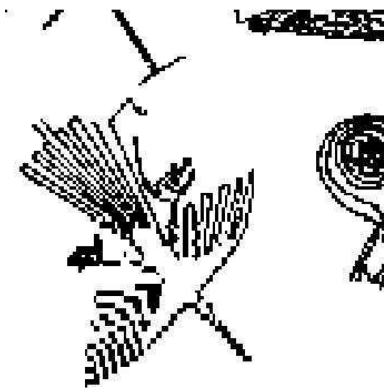
The authors declare that they have no competing interests.

References

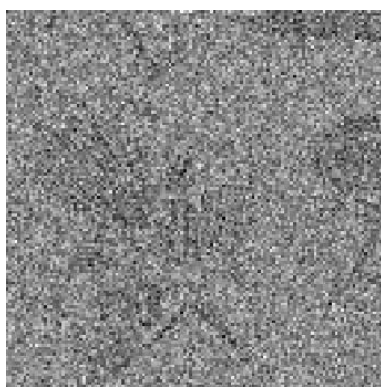
1. C. Raphael, Automatic segmentation of acoustic musical signals using hidden Markov models. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(4), 360–370 (1999)
2. M. Chen, A. Kundu, J. Zhou, Off-line handwritten word recognition using a hidden Markov model type stochastic network. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(5), 481–496 (1994)
3. O. Cappé, E. Moulines, T. Ryden, *Inference in Hidden Markov Models* (Springer, New York, 2005)
4. T. Koski, *Hidden Markov Models for Bioinformatics* (Kluwer Academic Publishers, Dordrecht, 2001)
5. L.C. Thomas, D.E. Allen, N. Morkel-Kingsbury, A hidden Markov chain model for the term structure of bond credit risk spreads. *Int. Rev. Financ. Anal.* **11**(3), 311–329 (2002)
6. F. Le Ber, M. Benoît, C. Schott, J.-F. Mari, C. Mignolet, Studying crop sequences with Carrot Age, a HMM based data mining software. *Ecol. Model.* **191**(1), 170–185 (2006)
7. L.E. Baum, T. Petrie, G. Soules, N. Weiss, A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.* **41**, 164–171 (1970)
8. G.D. Fornay, The Viterbi algorithm. *Proc. IEEE* **61**(3), 268–277 (1973)
9. L.R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* **77**(2), 257–286 (1989)
10. G. McLachlan, T. Krishnan, *EM Algorithm and Extensions* (Wiley, New York, 1997)
11. J.-P. Delmas, An equivalence of the EM and ICE algorithm for exponential family. *IEEE Trans. Signal Process.* **45**(10), 2613–2615 (1997)
12. P. Lanchantin, W. Pieczynski, Unsupervised restoration of hidden non stationary Markov chain using evidential priors. *IEEE Trans. Signal Process.* **53**(8), 3091–3098 (2005)
13. G. Shafer, *A Mathematical Theory of Evidence* (Princeton University Press, Princeton, 1976)

14. W. Pieczynski, Multisensor triplet Markov chains and theory of evidence. *Int. J. Approx. Reason.* **45**(1), 1–16 (2007)
15. W. Pieczynski, D. Benboudjema, Multisensor triplet Markov fields and theory of evidence. *Image Vis. Comput.* **24**(1), 61–69 (2006)
16. I. Bloch, Some aspects of Dempster–Shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account. *Pattern Recognit. Lett.* **17**(8), 905–919 (1996)
17. B.R. Cobb, P.P. Shenoy, *A Comparison of Methods for Transforming Belief Function Models to Probability Models*, vol. 2711 (Springer-Verlag, Berlin, 2003)
18. T. Denoeux, Analysis of evidence-theoretic decision rules for pattern classification. *Pattern Recognit.* **30**(7), 1095–1107 (1997)
19. L. Fouque, A. Appriou, W. Pieczynski, *An evidential Markovian model for data fusion and unsupervised image classification*, 3rd International Conference on Information Fusion (FUSION) (France, Paris, 2000), pp. 25–31
20. F. Janez, A. Appriou, Theory of evidence and non-exhaustive frames of discernment: plausibilities correction methods. *Int. J. Approx. Reason.* **18**(1), 1–19 (1998)
21. P. Smets, Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *Int. J. Approxim. Reason.* **9**(1), 1–35 (1993)
22. J.J. Sudano, *Equivalence between belief theories and naive Bayesian fusion for systems with independent evidential data*, Sixth International Conference on Information Fusion (FUSION) (Cairns, Australia, 2003), pp. 1357–1364
23. R.R. Yager, J. Kacprzyk, M. Fedrizzi, *Advances in the Dempster–Shafer Theory of Evidence* (Wiley, New York, 1994)
24. A. Bendjebbour, Y. Delignon, L. Fouque, V. Samson, W. Pieczynski, Multisensor images segmentation using Dempster–Shafer fusion in Markov fields context. *IEEE Trans. Geosci. Remote Sens.* **39**(8), 1789–1798 (2001)
25. P. Lanchantin, J. Lapuyade-Lahorgue, W. Pieczynski, Unsupervised segmentation of randomly switching data hidden with non-Gaussian correlated noise. *Signal Process.* **91**(2), 163–175 (2011)
26. S. Derrode, W. Pieczynski, Signal and image segmentation using Pairwise Markov chains. *IEEE Trans. Signal Process.* **52**(9), 2477–2489 (2004)
27. M.Y. Boudaren, W. Pieczynski, E. Monfrini, *Unsupervised segmentation of non-stationary data hidden with non-stationary noise*, IEEE International Workshop on Systems, Signal Processing and their Applications (WoSSPA) (Tipasa, Algeria, 2011), pp. 255–258
28. J. Lapuyade-Lahorgue, W. Pieczynski, Unsupervised segmentation of hidden semi-Markov non stationary chains. *Signal Process.* **92**(1), 29–42 (2012)

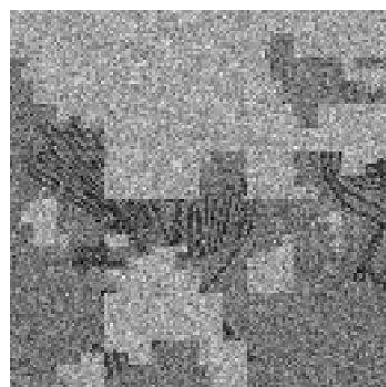
29. S.L. Hégarat-Masclé, I. Bloch, D. Vidal-Madjar, Introduction of neighborhood information on evidence theory and application to data fusion of radar and optical images with partial cloud cover. *Pattern Recognit.* **31**(11), 1811–1823 (1998)
30. W. Pieczynski, *EM and ICE in hidden and triplet Markov models*, *Stochastic Modeling Techniques and Data Analysis International Conference (SMTDA)* (Chania, Greece, 2010), pp. 1–9
31. R. Fjortoft, Y. Delignon, W. Pieczynski, M. Sigelle, F. Tupin, Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. *IEEE Trans. Geosci. Remote Sens.* **41**(3), 675–686 (2003)
32. A.S. Willsky, Multiresolution Markov models for signal and image processing. *Proc. IEEE* **90**(8), 1396–1458 (2002)



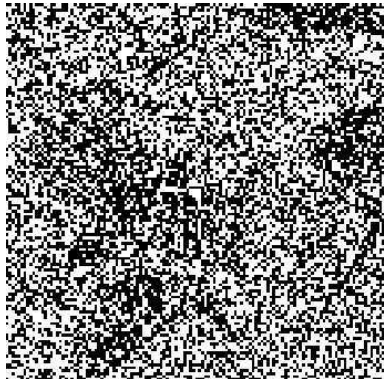
(a) $X = x$



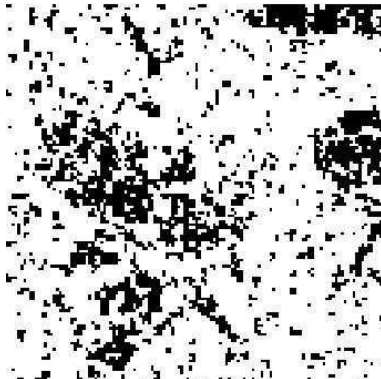
(b) $Y^1 = y^1$



(c) $Y^2 = y^2$



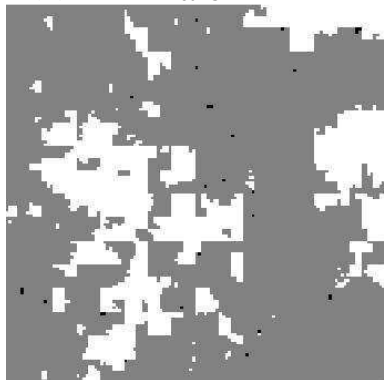
(d) $X = \hat{x}_{\text{K-Means}}, \tau = 39.4\%$



(e) $X = \hat{x}_{\text{HMC}}, \tau = 16.5\%$



(f) $X = \hat{x}_{\text{EHMC}}, \tau = 11.1\%$



(g) $U^1 = \hat{u}_{\text{EHMC}}^1$



(h) $X = \hat{x}_{\text{MS-HMC}}, \tau = 12\%$



(i) $U^2 = \hat{u}_{\text{MS-HMC}}^2, \tau = 0.1\%$



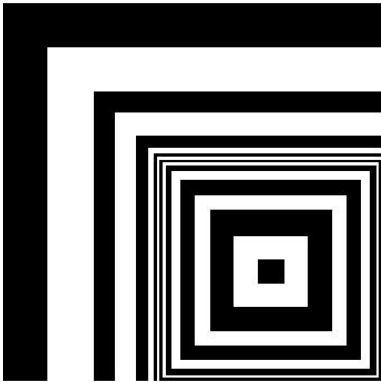
(j) $X = \hat{x}_{\text{MN-HMC}}, \tau = 7.9\%$



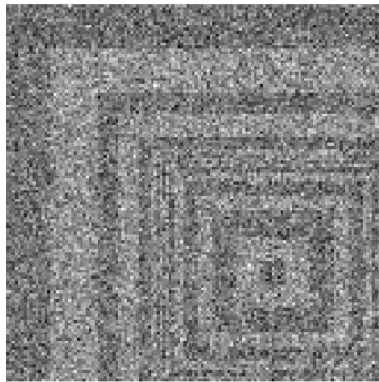
(k) $U^1 = \hat{u}_{\text{MN-HMC}}^1$



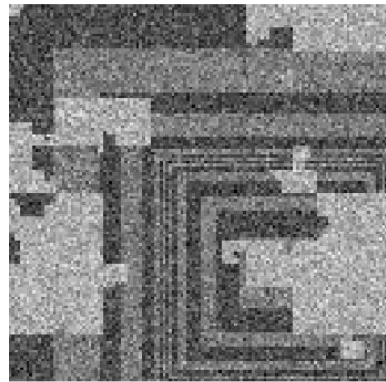
(l) $U^2 = \hat{u}_{\text{MN-HMC}}^2, \tau = 0.1\%$



(a) $X = x$



(b) $Y^1 = y^1$



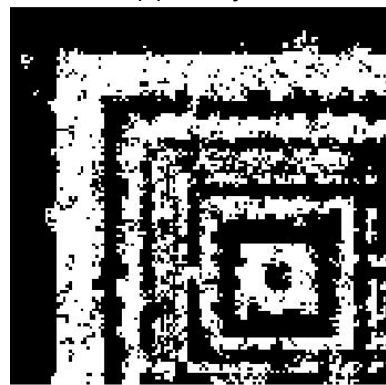
(c) $Y^2 = y^2$



(d) $X = \hat{x}_{K\text{-Means}}, \tau = 30.9\%$



(e) $X = \hat{x}_{\text{HMC}}, \tau = 16.4\%$



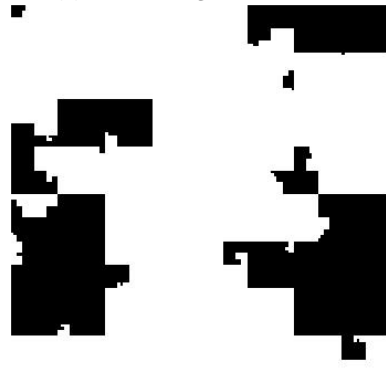
(f) $X = \hat{x}_{\text{EHMC}}, \tau = 15.9\%$



(g) $U^1 = \hat{u}_{\text{EHMC}}^1$



(h) $X = \hat{x}_{\text{MS-HMC}}, \tau = 9.1\%$



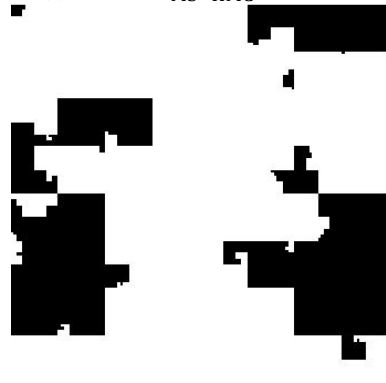
(i) $U^2 = \hat{u}_{\text{MS-HMC}}^2, \tau \approx 0\%$



(j) $X = \hat{x}_{\text{MN-HMC}}, \tau = 7.6\%$



(k) $U^1 = \hat{u}_{\text{MN-HMC}}^1$



(l) $U^2 = \hat{u}_{\text{MN-HMC}}^2, \tau \approx 0\%$